Covid case analysis Project Phase 3 - Documentation

Phase 3: Development Part 1

In Phase 3, we will begin building the Covid case analysis project. This phase focuses on loading and preprocessing the dataset, defining analysis objectives, and using IBM Cognos for data visualization.

Step 1: Dataset Download

- Access the dataset from the provided Kaggle link: [Kaggle Covid 19 Dataset](https://www.kaggle.com/datasets/imdevskp/corona-virus-report).
- Download the dataset to your local working directory or preferred location.

Step 2: Loading the Dataset

- Import the necessary Python libraries, including Pandas, for covid case analysis.
- Load the dataset into a Pandas DataFrame for further analysis.
- Display the first few rows of the dataset to inspect the data structure.

Step 3: Exploratory Data Analysis

- Perform an initial exploration of the dataset to understand the observation
- Check for basic statistics and disease spread rate
- Visualize key features to gain insights into the data.
- Identify potential relationships or correlations between the year/date/month and death rate

Step 4: Define Analysis Objectives

Define the objectives of the analysis for this phase, including: • Covid Case analysis

Prediction: Develop a model based on the covid case observation.

- Death rate: Number of death occurred based on the covid spread
- Year/Date/Month: spread of diease and death rate is measured based on this

Step 5: Data Cleaning and Preprocessing

Clean and preprocess the data to ensure its quality and suitability for analysis:

- Integrate the data from the source
- Aggregate it into weekly or monthly record
- Verify the correctness of the integrated data

Step 6: IBM Cognos for Visualization

Utilize IBM Cognos for creating data visualizations that provide insights into the dataset:

- Create various visualizations, such as bar charts, line charts, heatmaps, and interactive dashboards.
- Visualize the rate of death of covid outspread,feature ,other relevant insights using IBM Cognos.

Step 7: Data Validation

Validate the processed data to ensure its quality and accuracy:

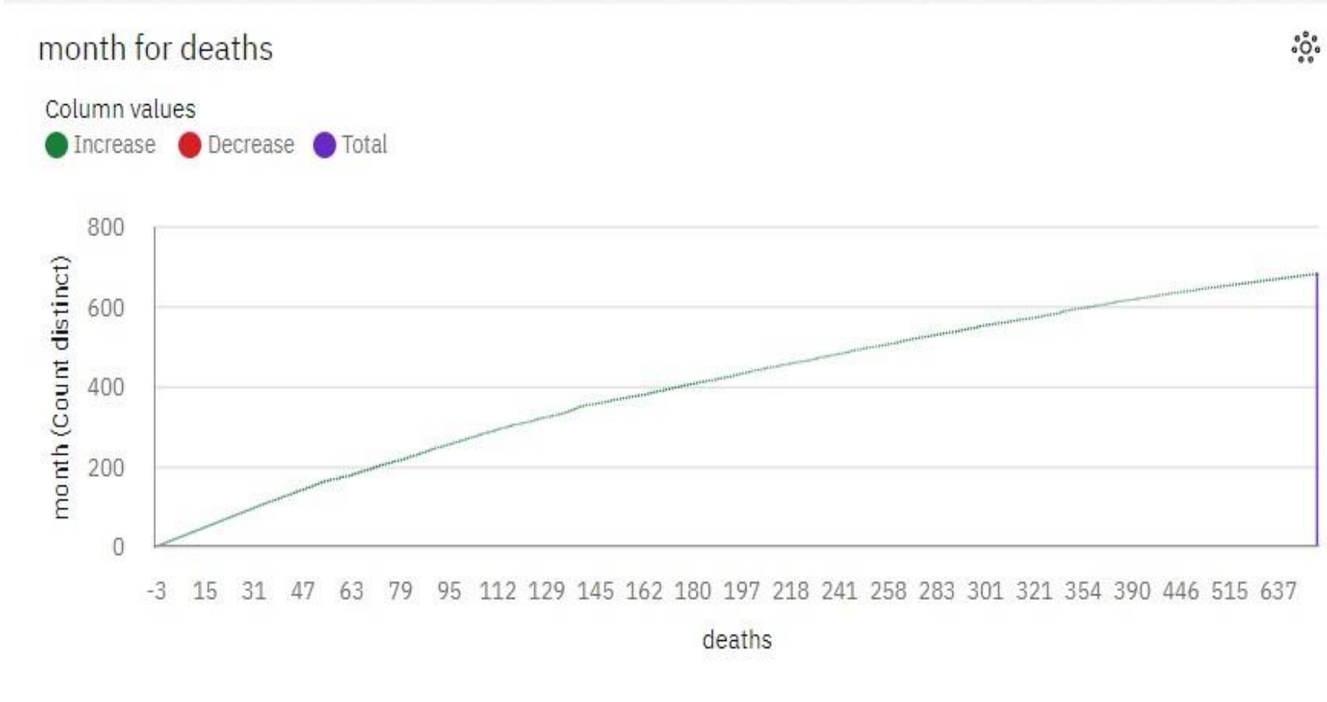
- Perform data validation checks to confirm data consistency and accuracy.
- Identify and address any potential data quality issues.

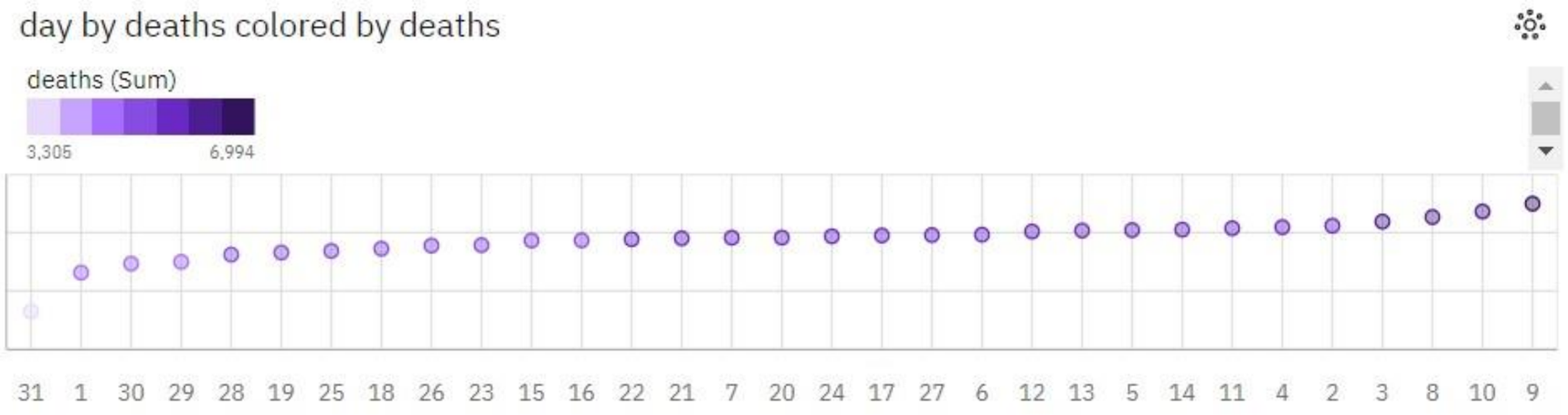
Step 8: Documentation

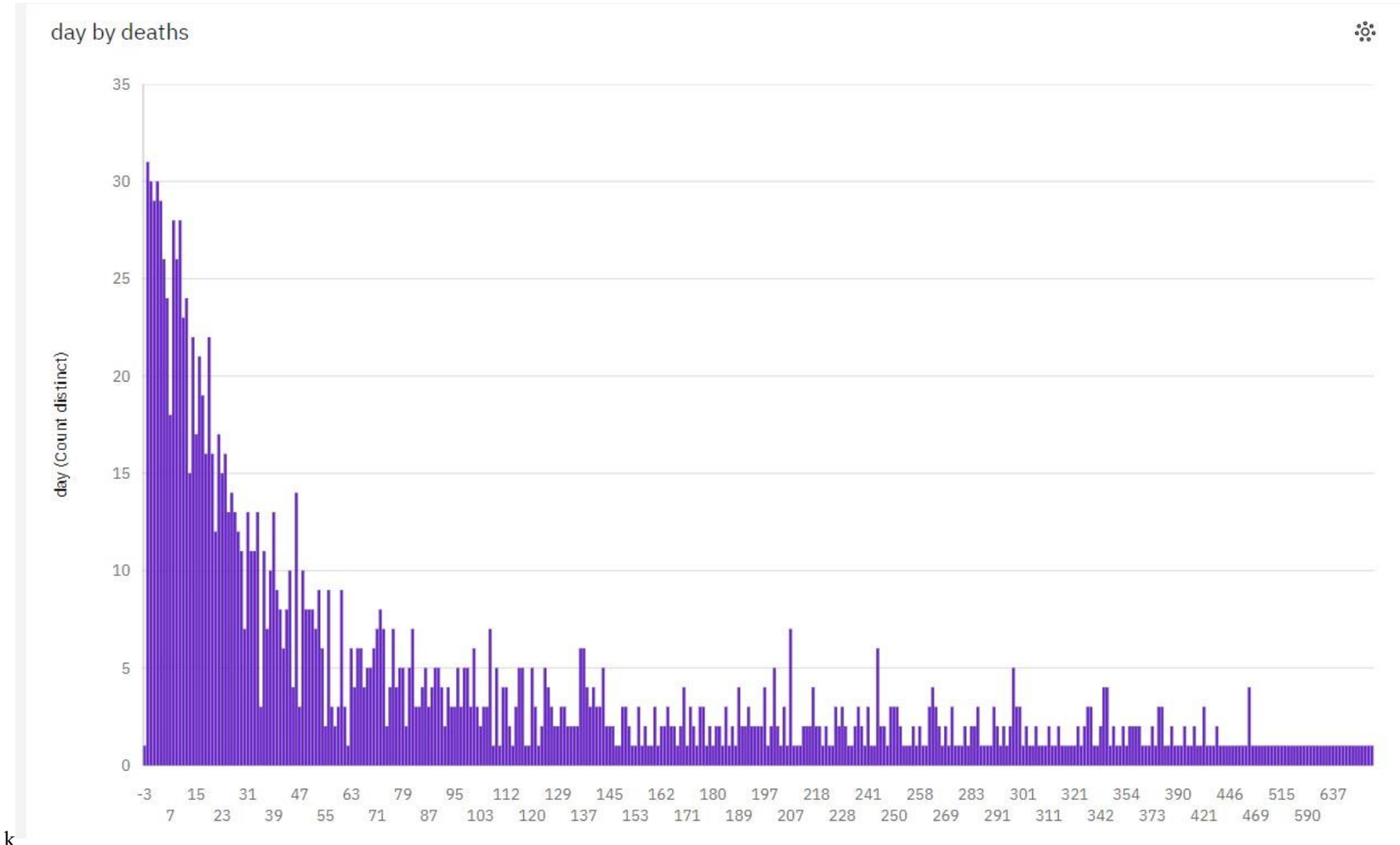
Maintain comprehensive documentation of the activities performed in this phase:

- Record the data preprocessing steps, including any transformations and cleaning.
- Document the objectives defined for this phase.
- Capture insights gained from data visualization using IBM Cognos.
- Document any observations, challenges, or discoveries made during this phase.

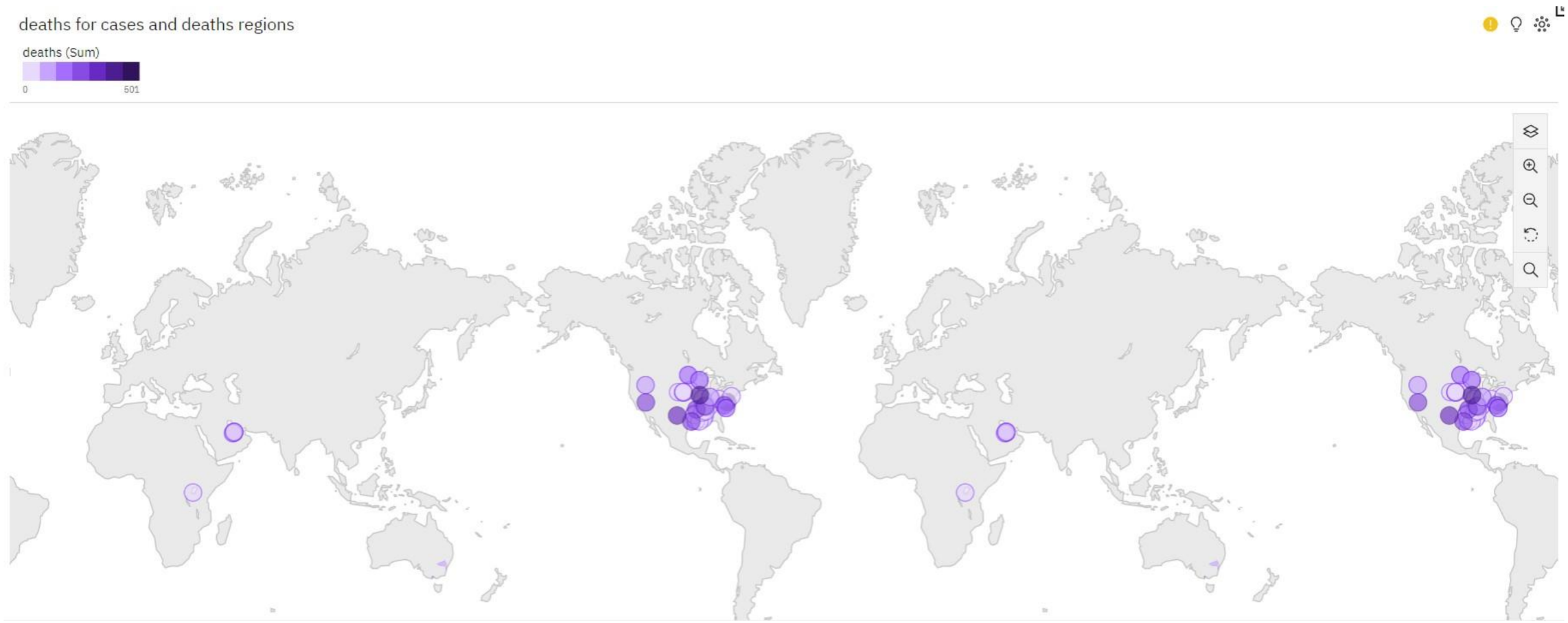
This documentation outlines the key steps and activities to be performed in Phase 3 of the Covid case analysis project. It serves as a guide for the project's development, dataset preparation, and the use of IBM Cognos for data visualization.

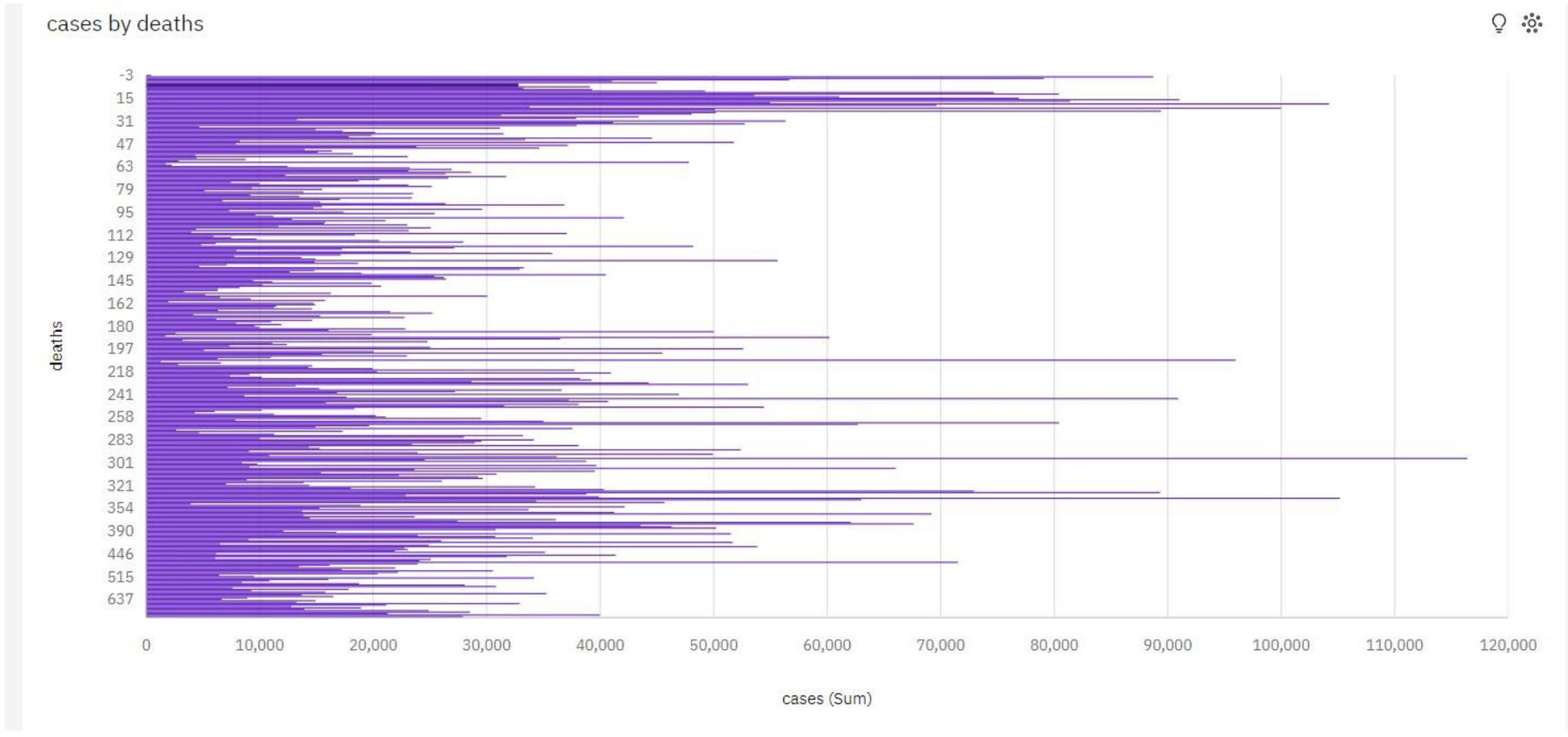

month for deaths
Column values
Increase
Decrease
Total
800
600
400
200
0
month (Count distinct)
-3 15 31 47 63 79 95 112 129 145 162 180 197 218 241 258 283 301 321 354 390 446 515 637
deaths


day by deaths colored by deaths
deaths (Sum)
3,305
6,994
31 1 30 29 28 19 25 18 26 23 15 16 22 21 7 20 24 17 27 6 12 13 5 14 11 4 2 3 8 10 9


day by deaths
day (Count distinct)
35
30
25
20
15
10
5
0
-3 15 31 47 63 79 95 112 129 145 162 180 197 218 241 258 283 301 321 354 390 446 515 637
7 23 39 55 71 87 103 120 137 153 171 189 207 228 250 269 291 311 342 373 421 469 590

k


deaths for cases and deaths regions
deaths (Sum)
0
501


cases by deaths
deaths
-3
15
31
47
63
79
95
112
129
145
162
180
197
218
241
258
283
301
321
354
390
446
515
637
0
10,000
20,000
30,000
40,000
50,000
60,000
70,000
80,000
90,000
100,000
110,000
120,000
cases (Sum)